Wavelet-Domain Compressive Signal Reconstruction Using a Hidden Markov Tree Model

Marco F. Duarte

RICE

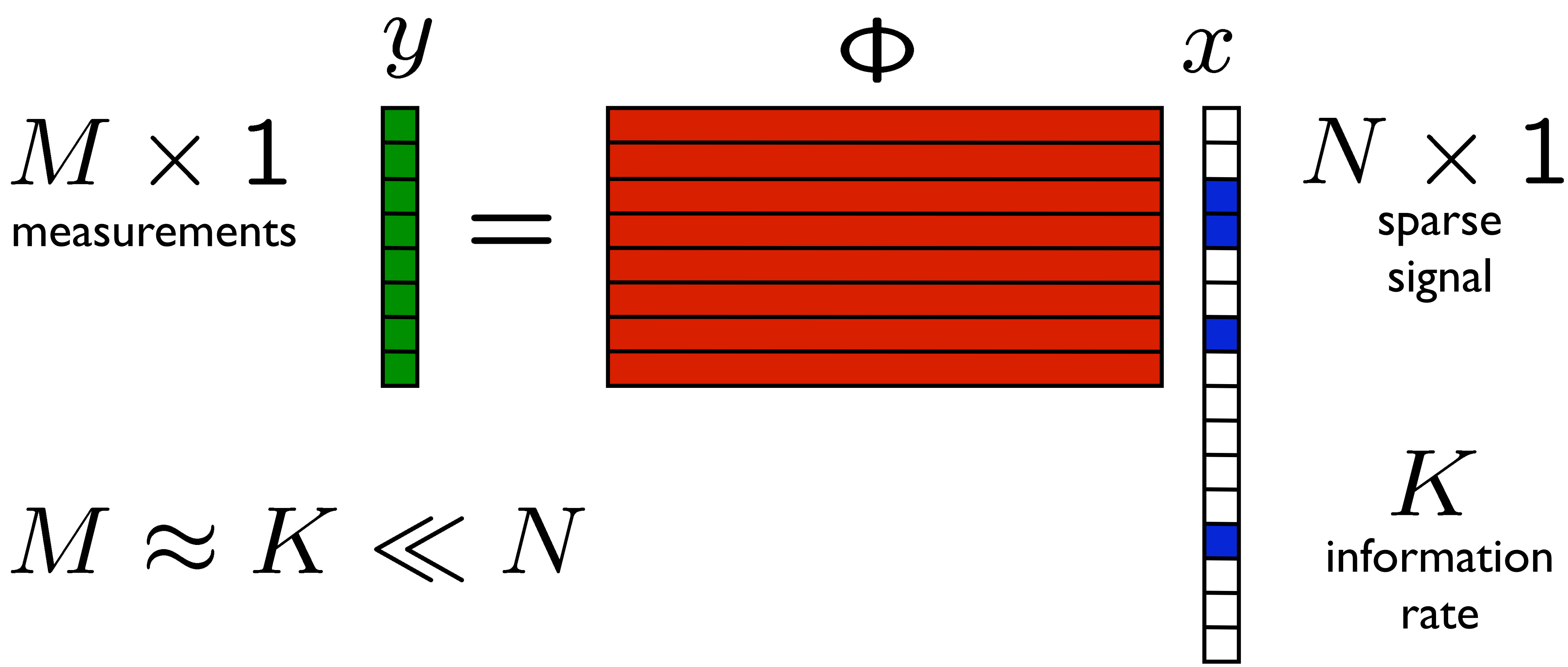
Joint work with M. B. Wakin and R. G. Baraniuk
ICASSP - April 1, 2008

THE UNIVERSITY OF MICHIGAN
ARTES · SCIENTIA · VERITAS
1817

Compressive Sensing:
From Samples to *Measurements*

- Many interesting signals are ***sparse*** or ***compressible***
- Instead of **sampling** the signal**,** ***encode*** the relevant information into a few measurements (inner products)

$$y = \Phi x$$

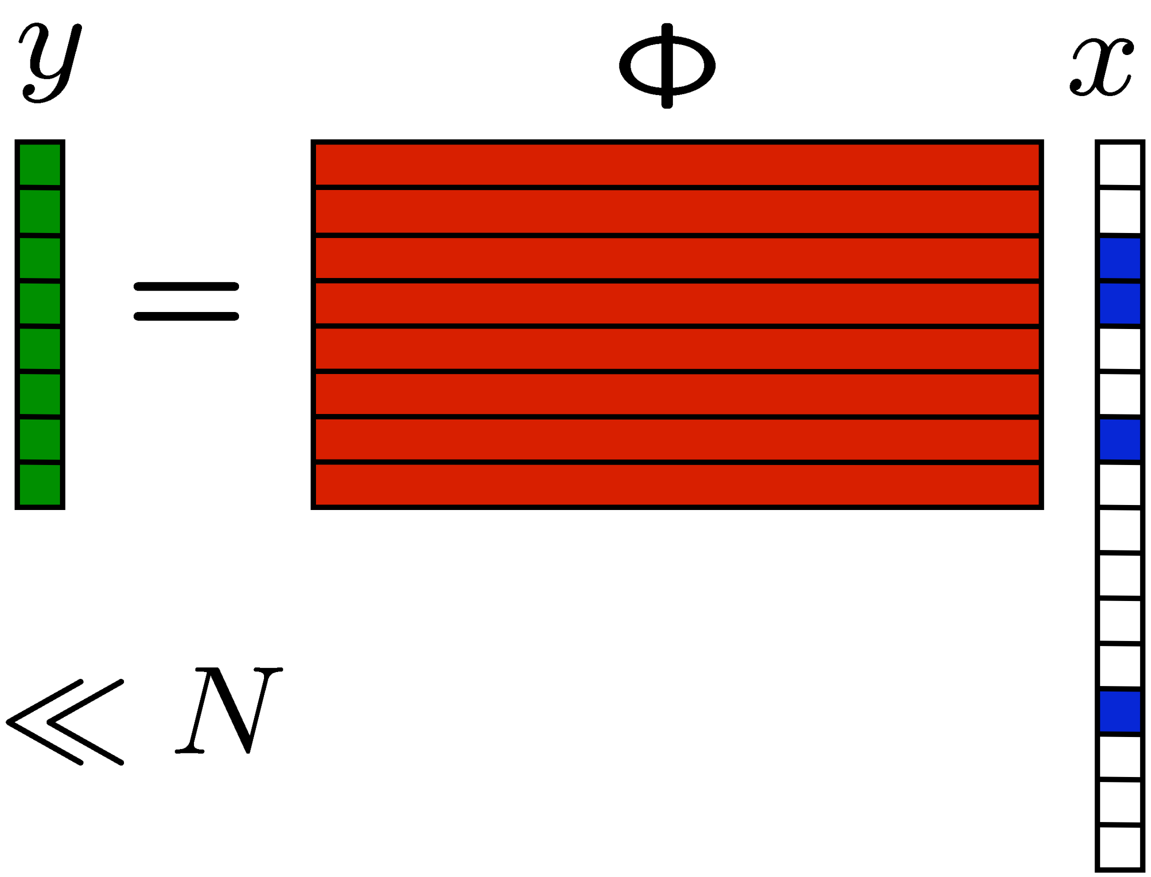


$$M \approx K \ll N$$

Compressive Sensing:
From Samples to *Measurements*

- **Decoding** = reconstruction = *inverse problem*
  - given y, extract information of interest about x

$$y = \Phi x$$



$M \times 1$
measurements

$N \times 1$
sparse
signal

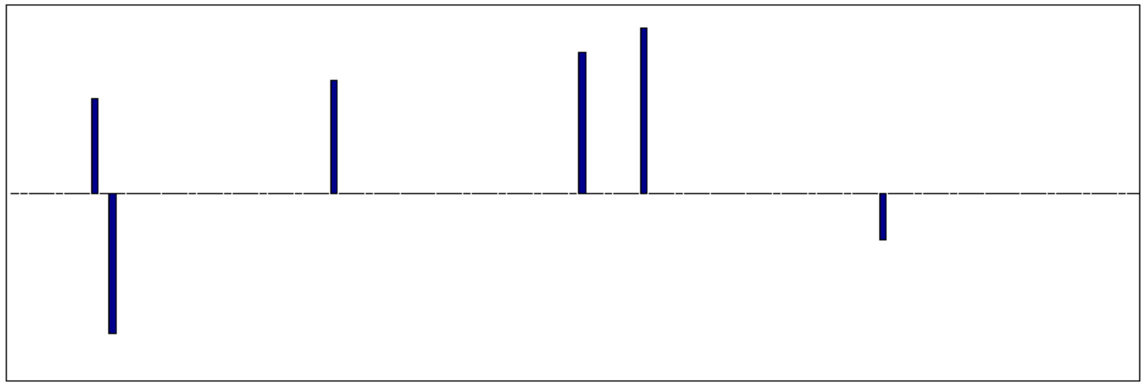
K
information
rate

$$M \approx K \ll N$$

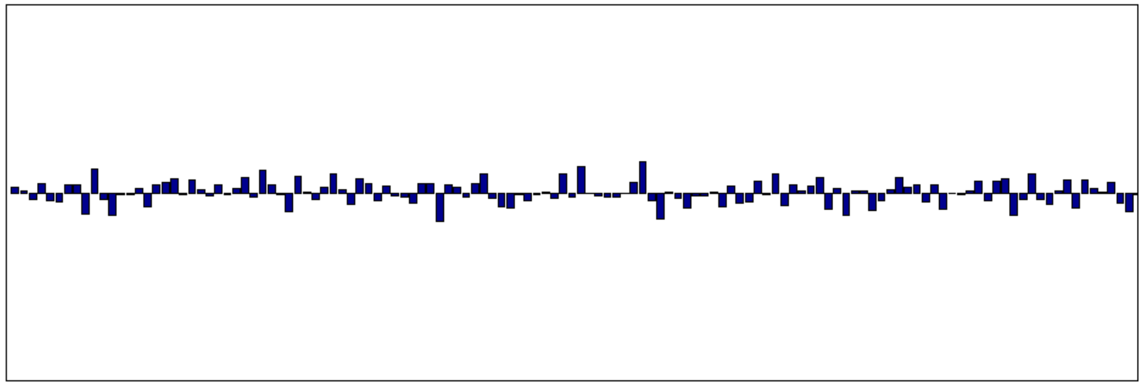
CS Signal Recovery

- Reconstruction/decoding: given $y = \Phi x$
  (ill-posed inverse problem) find x

- ℓ_2 fast, **wrong** $\widehat{x} = \arg\min_{y=\Phi x} \|x\|_2$



x



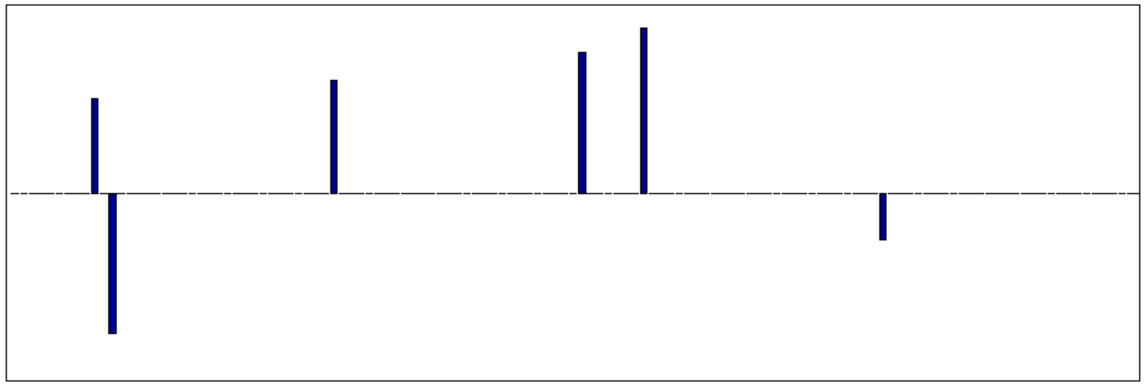
$\widehat{x} = (\Phi^T \Phi)^{-1} \Phi^T y$

CS Signal Recovery

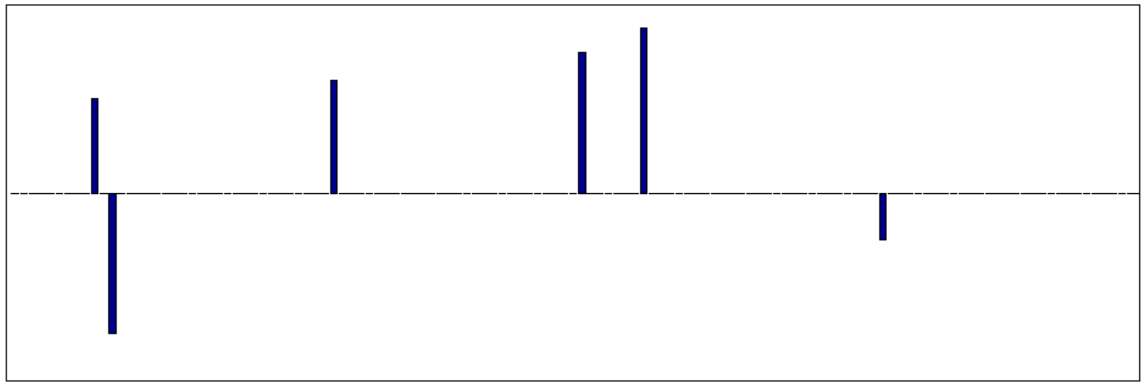
- Reconstruction/decoding: given $y = \Phi x$
  (ill-posed inverse problem) find x

- ℓ_2 fast, wrong $\quad \widehat{x} = \arg\min_{y=\Phi x} \|x\|_2$

- ℓ_0 **correct,** $M = 2K$ **slow** $\quad \widehat{x} = \arg\min_{y=\Phi x} \|x\|_0$



x



$\widehat{x}$

CS Signal Recovery

- Reconstruction/decoding: given $y = \Phi x$
  (ill-posed inverse problem) find x

- ℓ_2 fast, wrong $\widehat{x} = \text{arg}\min_{y=\Phi x} \|x\|_2$

- ℓ_0 correct, $M = 2K$
  slow $\widehat{x} = \text{arg}\min_{y=\Phi x} \|x\|_0$

- ℓ_1 **correct,**
  $M = CK \log(N/K)$
  [Candès et al, Donoho, ...] $\widehat{x} = \text{arg}\min_{y=\Phi x} \|x\|_1$
  linear program

- greedy [Tropp, Gilbert, Strauss; Rice]
- statistical [Candès; Nowak et al; Rice]

Reweighted ℓ_1 Algorithm

- Minimize $\|Wx\|_1 = \sum_n w_n |x_n|$
- Use iterative procedure to construct weights

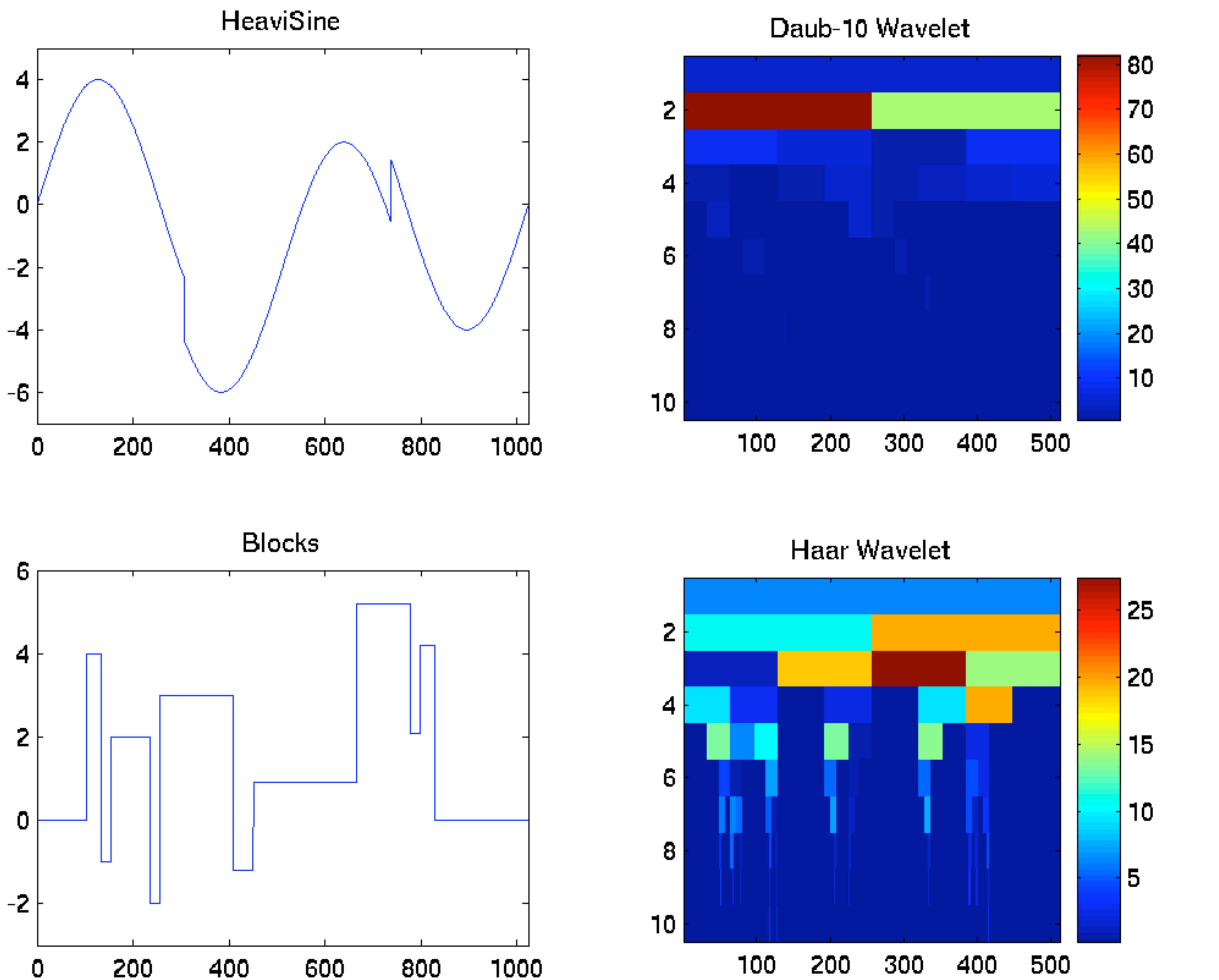
$$w_n^{(0)} = 1 \text{ for all } n$$
$$\widehat{x}^{(i)} = \arg\min_{y=\Phi x} \|W^{(i-1)} x\|_1$$
$$w_n^{(i)} = \frac{1}{|\widehat{x}_n^{(i)}| + \epsilon} \text{ for all } n$$

so that $\|Wx\|_1 \approx \|x\|_0$

- Weights w_n small when $|x_n|$ large
- Reduce number of measurements M required for reconstruction

[Candès, Wakin and Boyd. See also Gorodnitsky and Rao; Figueiredo, Bioucas-Dias, and Nowak]

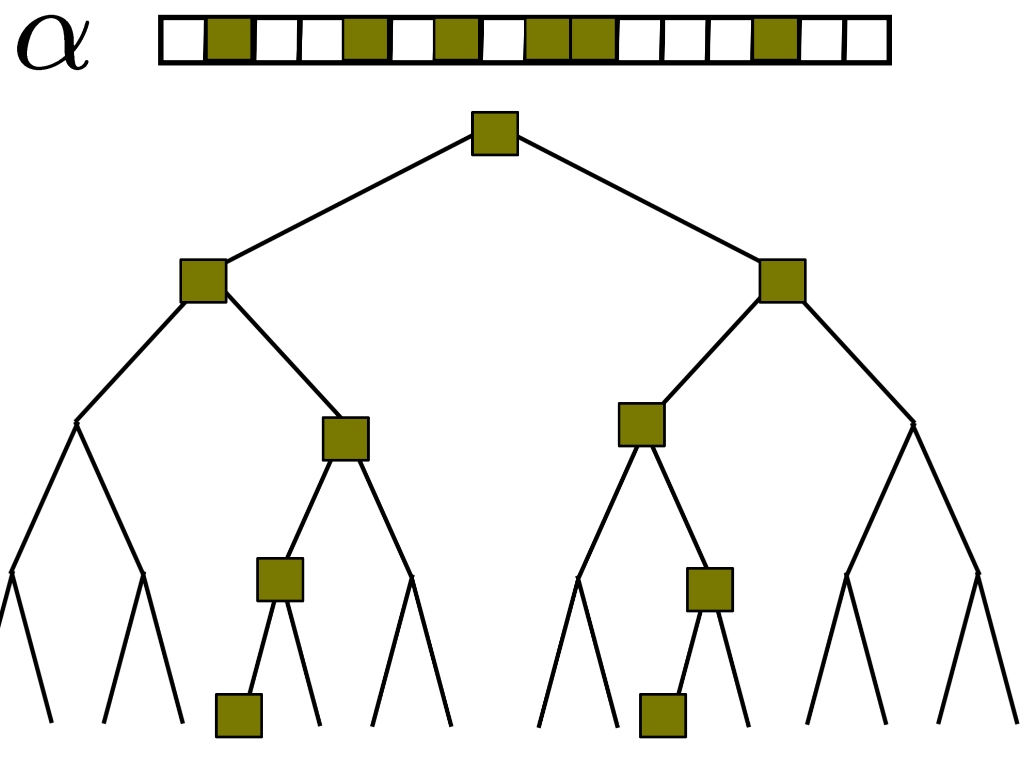
Wavelets: more than just sparsity



Wavelets: more than just sparsity

- Wavelet transform sports a *connected subtree structure*
- Piecewise smooth signal "rule of thumb"
  - *persistence*: small/large values tumble down the tree
  - *magnitude*: wavelet coefficients decay monotonically along branches of wavelet tree
- **Exploit this structure** for
  - fast reconstruction
  - lower oversampling
  - noise regularization

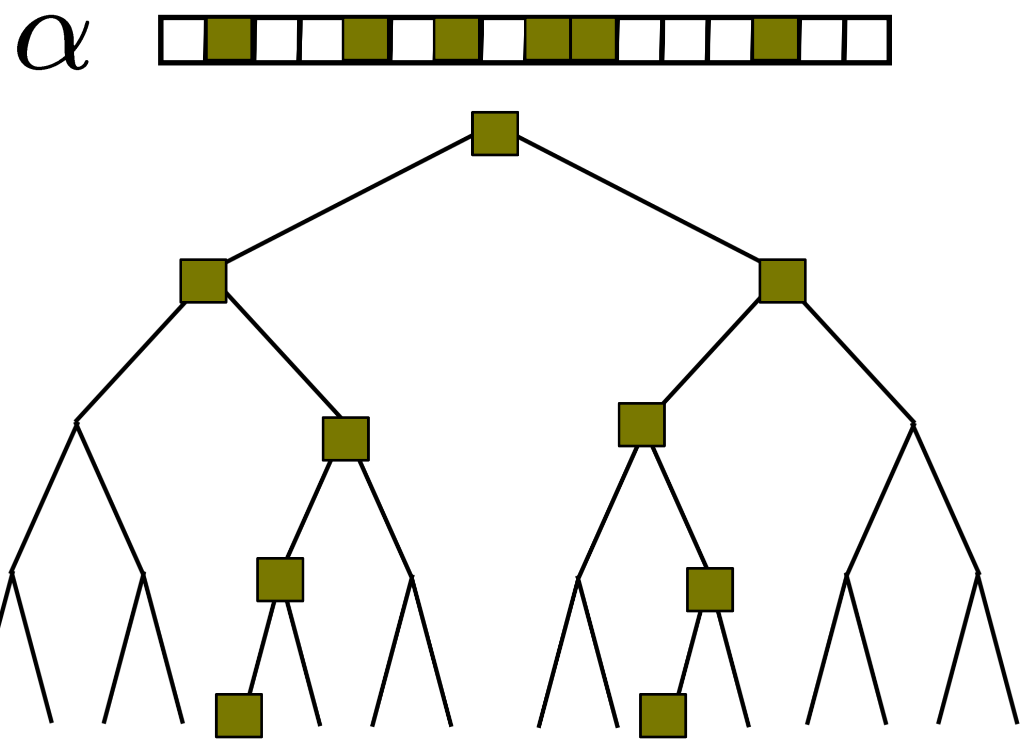
$$x = \Psi\alpha$$



Wavelets: more than just sparsity

- Wavelet transform sports a *connected subtree structure*
- Piecewise smooth signal "rule of thumb"
  - *persistence*: small/large values tumble down the tree
  - *magnitude*: wavelet coefficients decay monotonically along branches of wavelet tree
- **Exploit this structure** for
  - fast reconstruction
  - lower oversampling
  - noise regularization
- **Greedy Algorithms**
  - restrict search to connected tree [Duarte/Wakin/Baraniuk, La/Do]

$$x = \Psi\alpha$$

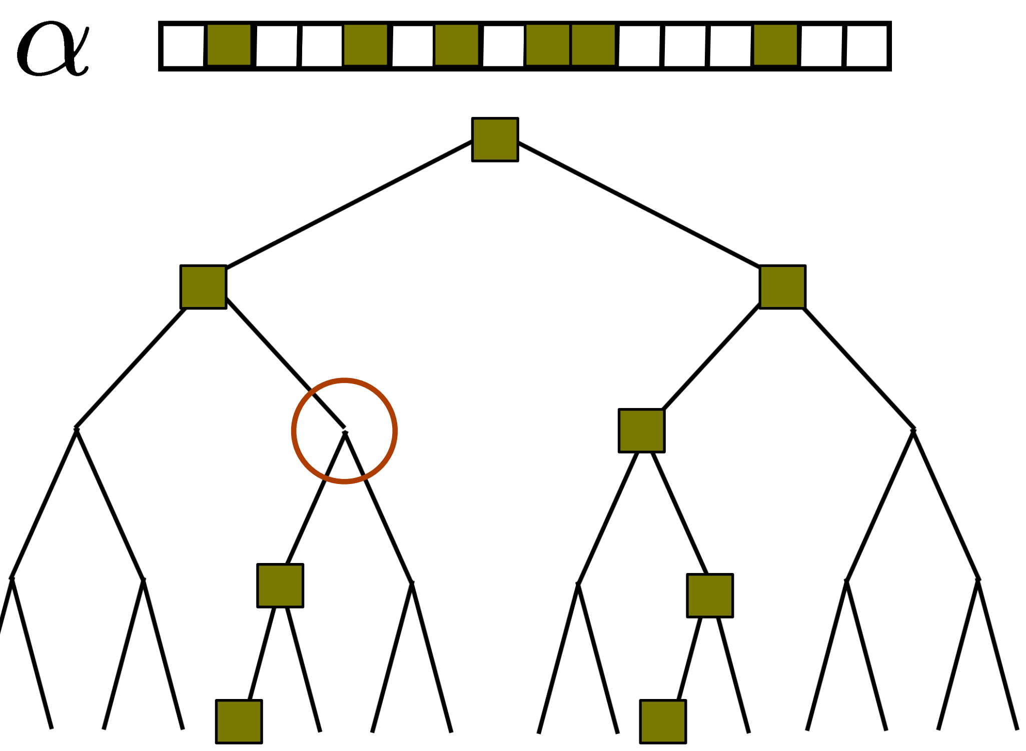


Wavelets: more than just sparsity

- Wavelet transform sports a *connected subtree structure*

$$x = \Psi\alpha$$

- Piecewise smooth signal "rule of thumb"
  - *persistence*: small/large values tumble down the tree
  - *magnitude*: wavelet coefficients decay monotonically along branches of wavelet tree



- **Exploit this structure** for
  - fast reconstruction
  - lower oversampling
  - noise regularization
- **Greedy Algorithms**
  - restrict search to connected tree [Duarte/Wakin/Baraniuk, La/Do]
- **Issues**: Gaps in branches of large-coefficient subtree

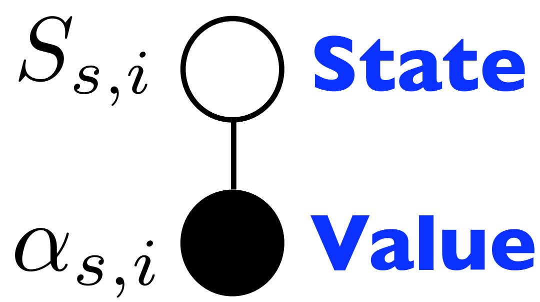
Hidden Markov Tree

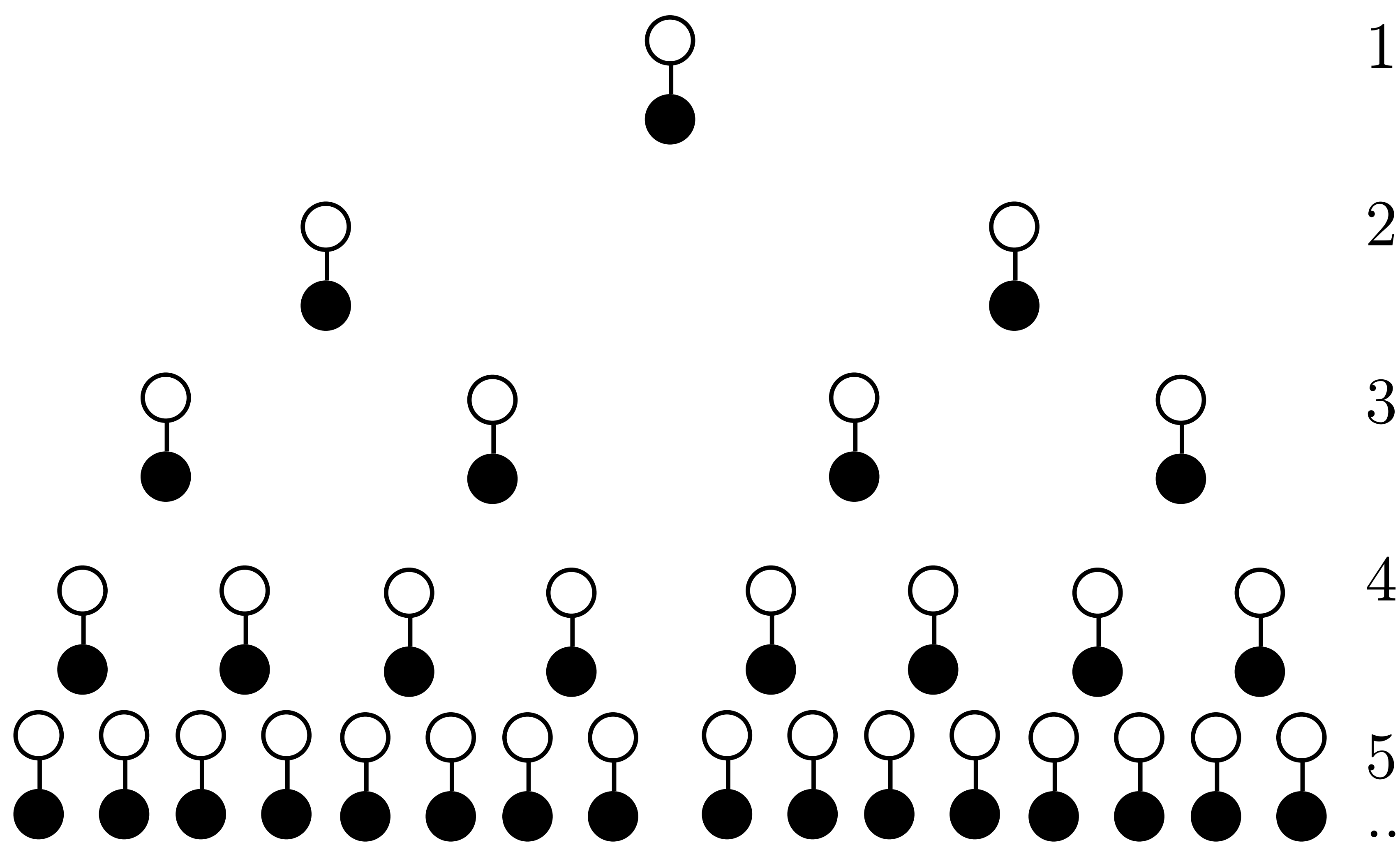
http://www.flickr.com/photos/bestrated1/

Hidden Markov Tree

http://www.flickr.com/photos/bestrated1/

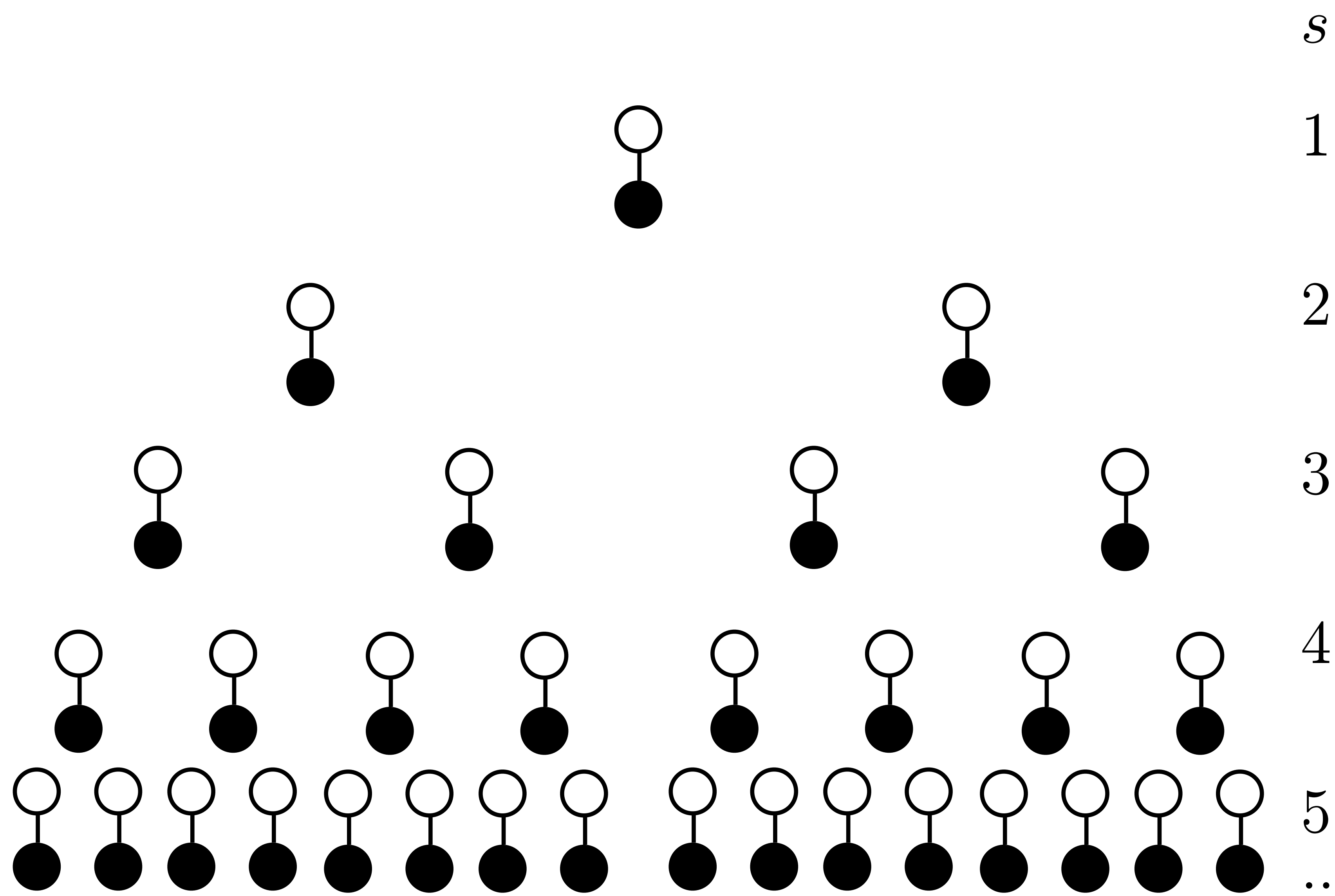
Hidden Markov Tree (HMT)





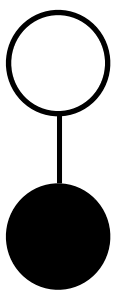
[Crouse, Nowak and Baraniuk, 1997]

Hidden Markov Tree (HMT)



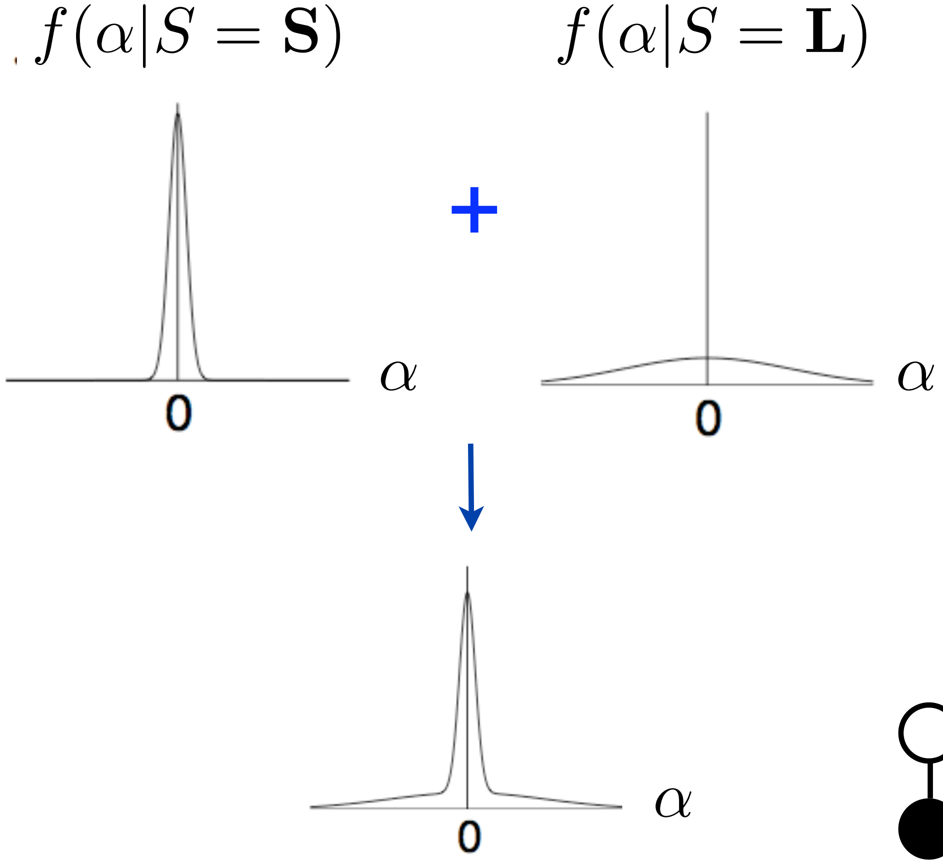
[Crouse, Nowak and Baraniuk, 1997]

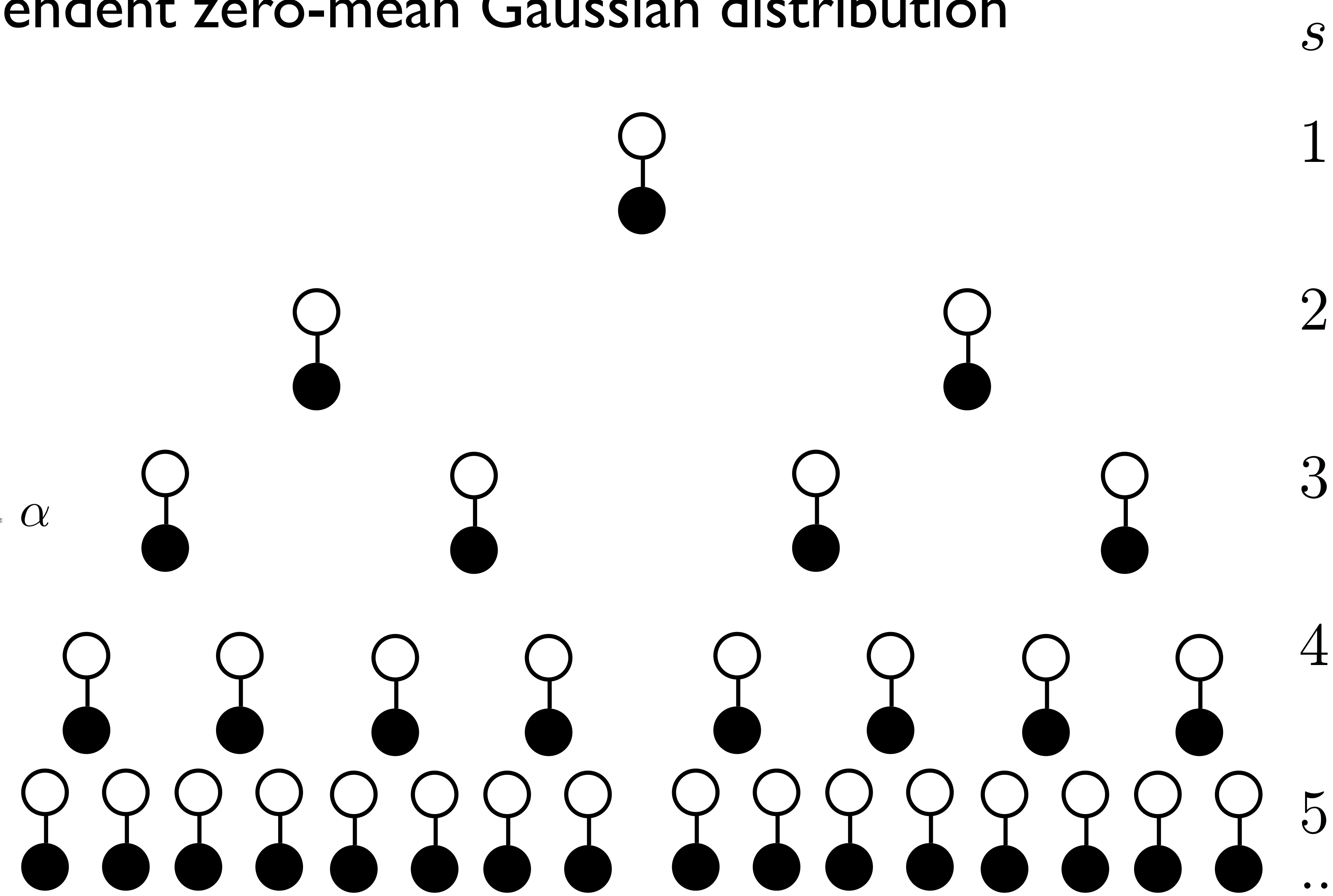
Hidden Markov Tree (HMT)



$S_{s,i}$ **State:** Large, Small

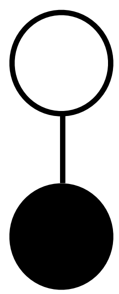
$\alpha_{s,i}$ **Value:** State-dependent zero-mean Gaussian distribution





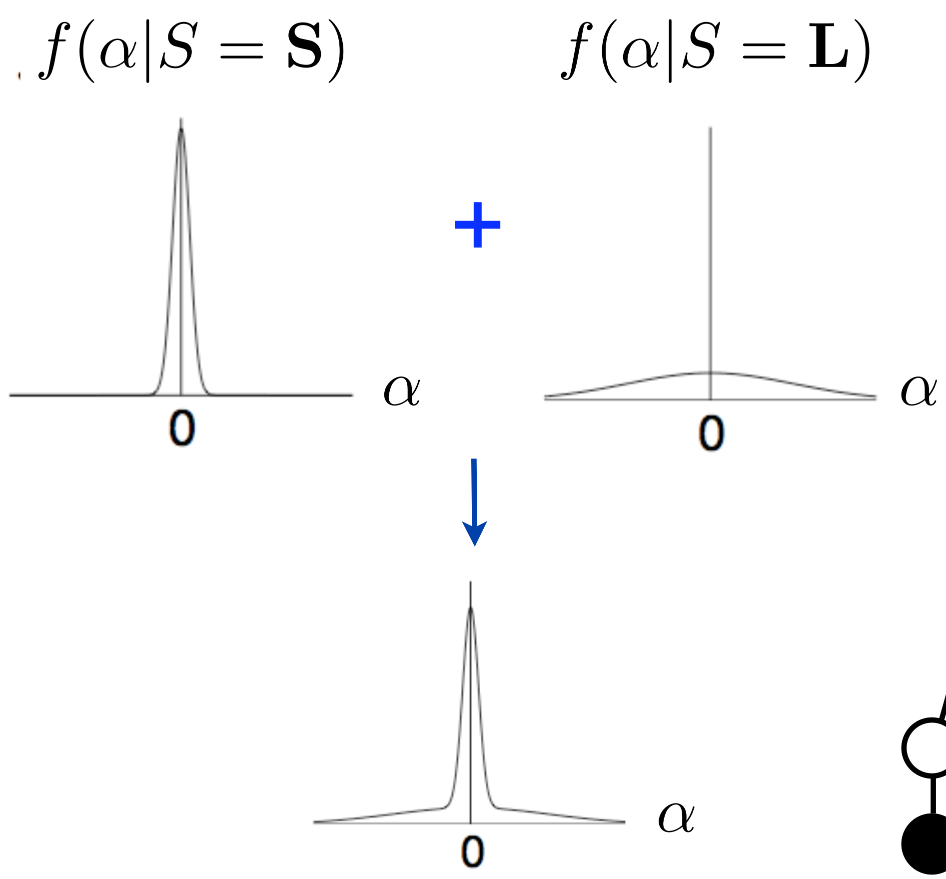
[Crouse, Nowak and Baraniuk, 1997]

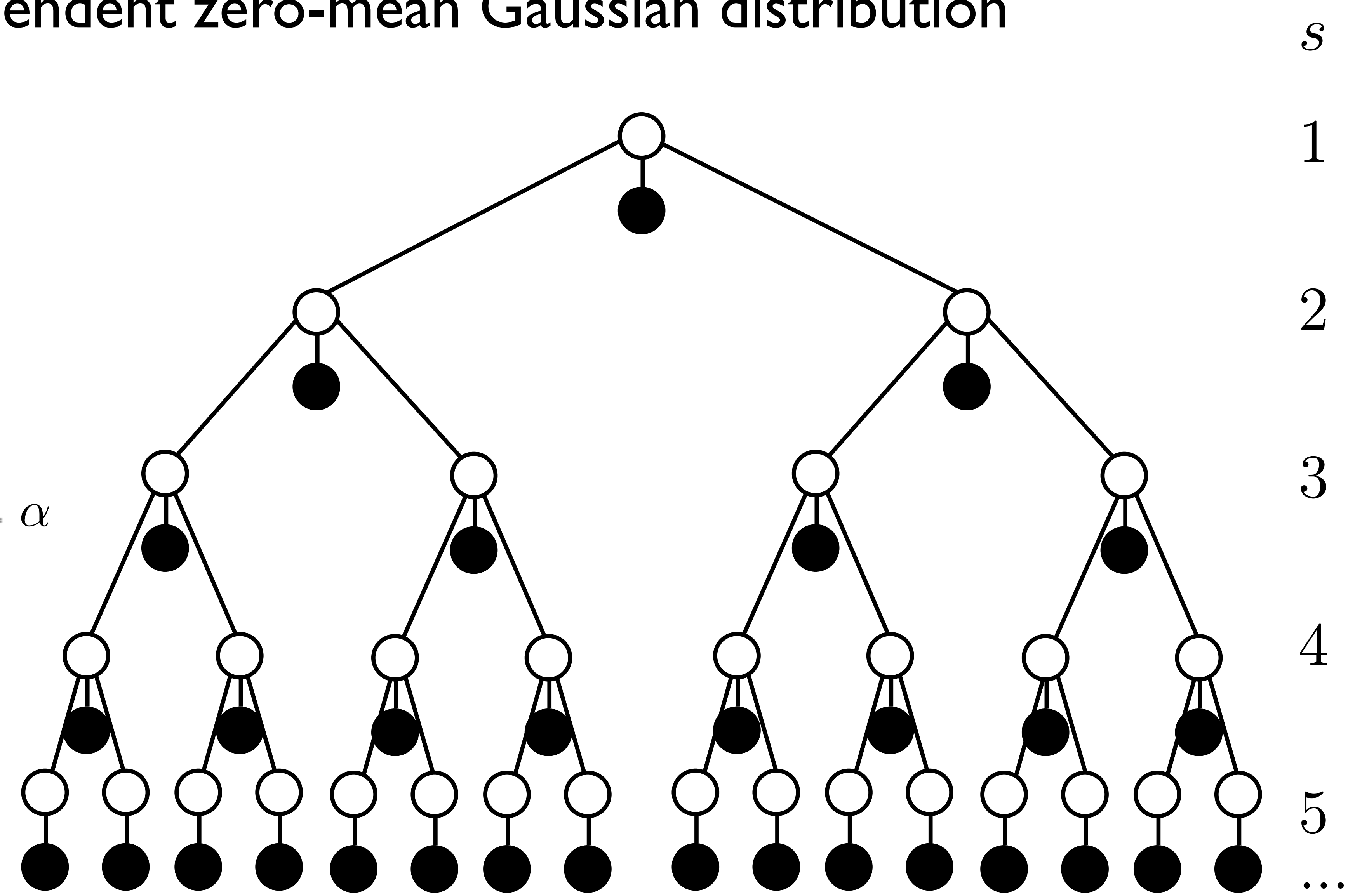
Hidden Markov Tree (HMT)



$S_{s,i}$ **State:** Large, Small

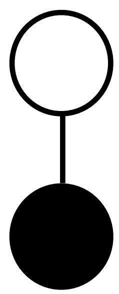
$\alpha_{s,i}$ **Value:** State-dependent zero-mean Gaussian distribution





[Crouse, Nowak and Baraniuk, 1997]

Hidden Markov Tree (HMT)

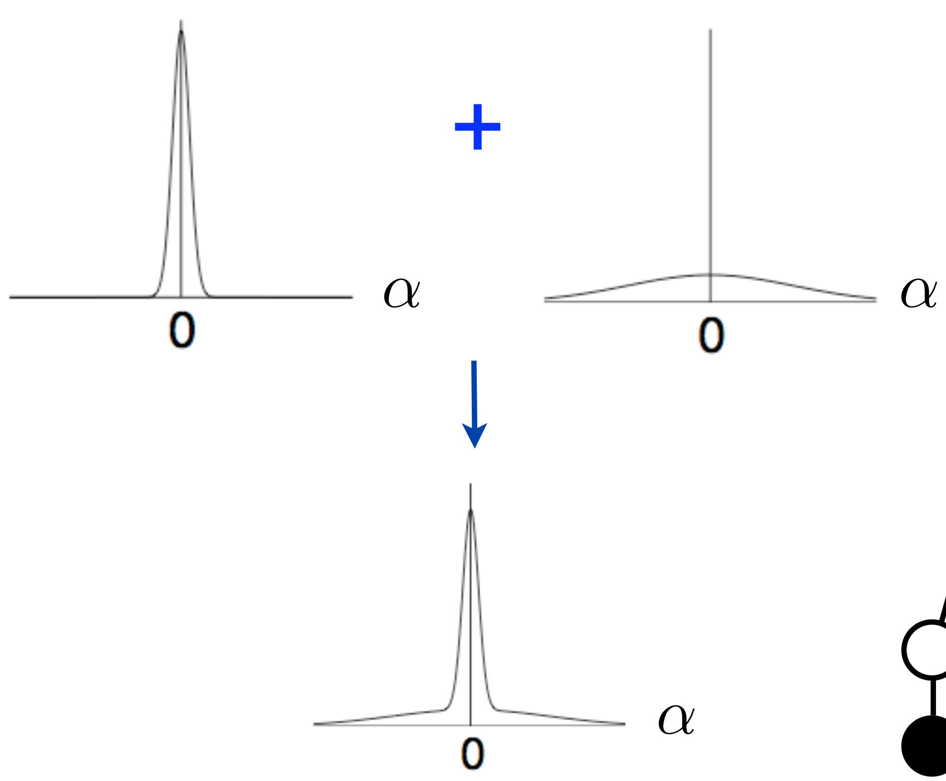


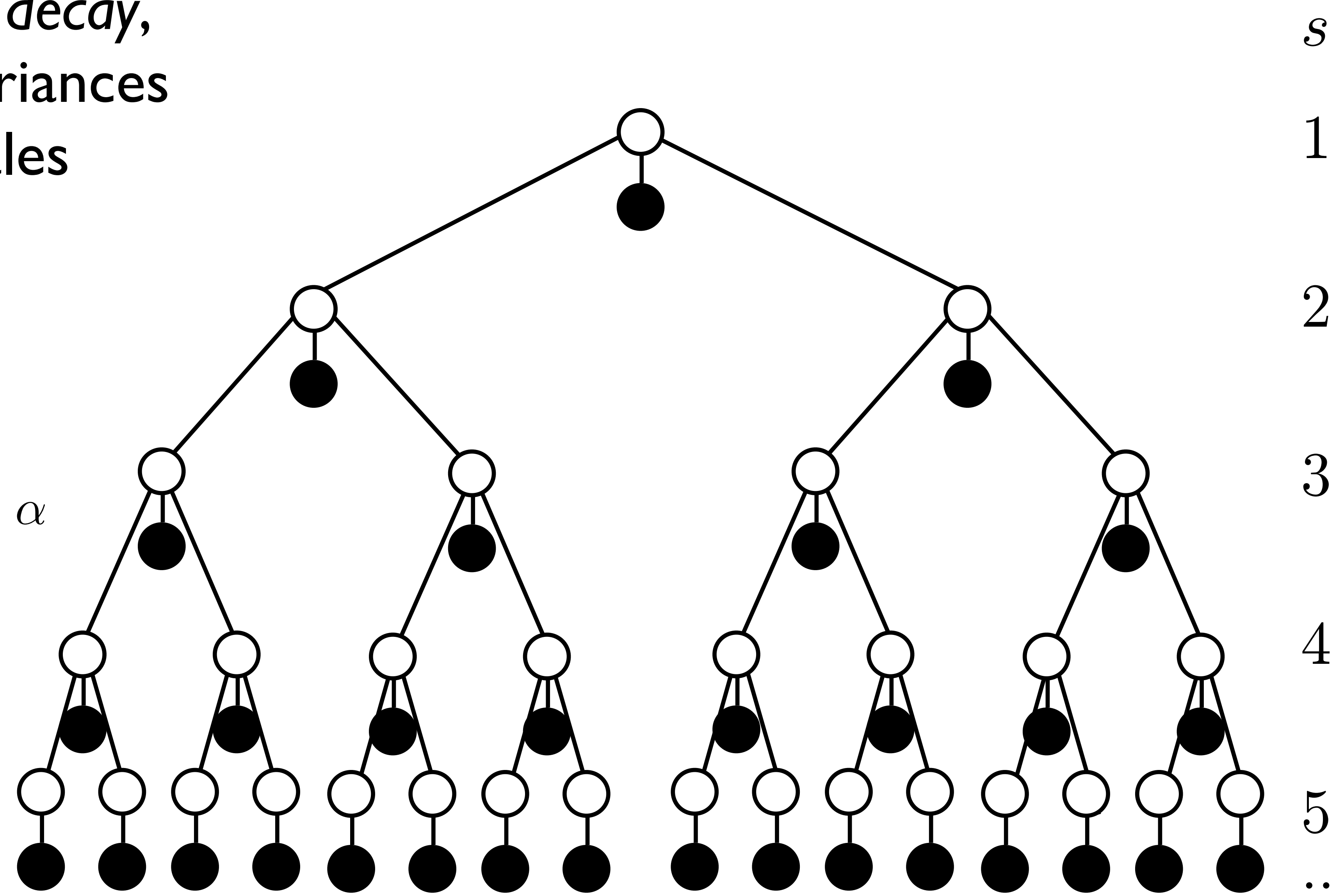
$S_{s,i}$ **State:** To obtain *persistence*, favor progressions

$\alpha_{s,i}$ **Value:** To obtain *decay*, reduce variances across scales

$$\mathbf{L} \rightarrow \mathbf{L} \rightarrow \mathbf{L} \rightarrow \ldots$$
$$\mathbf{S} \rightarrow \mathbf{S} \rightarrow \mathbf{S} \rightarrow \ldots$$

$f(\alpha|S = \mathbf{S})$ $f(\alpha|S = \mathbf{L})$

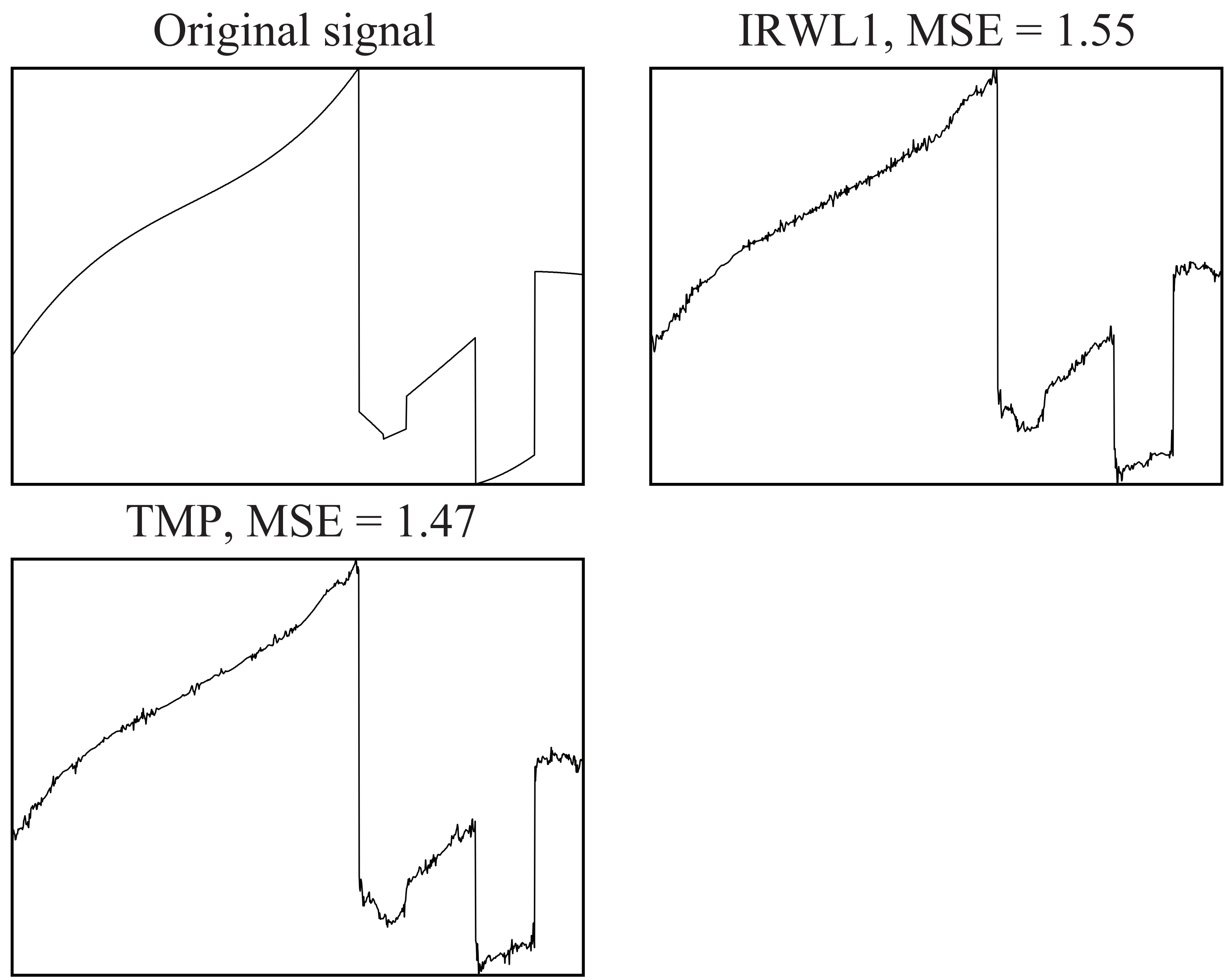




[Crouse, Nowak and Baraniuk, 1997]

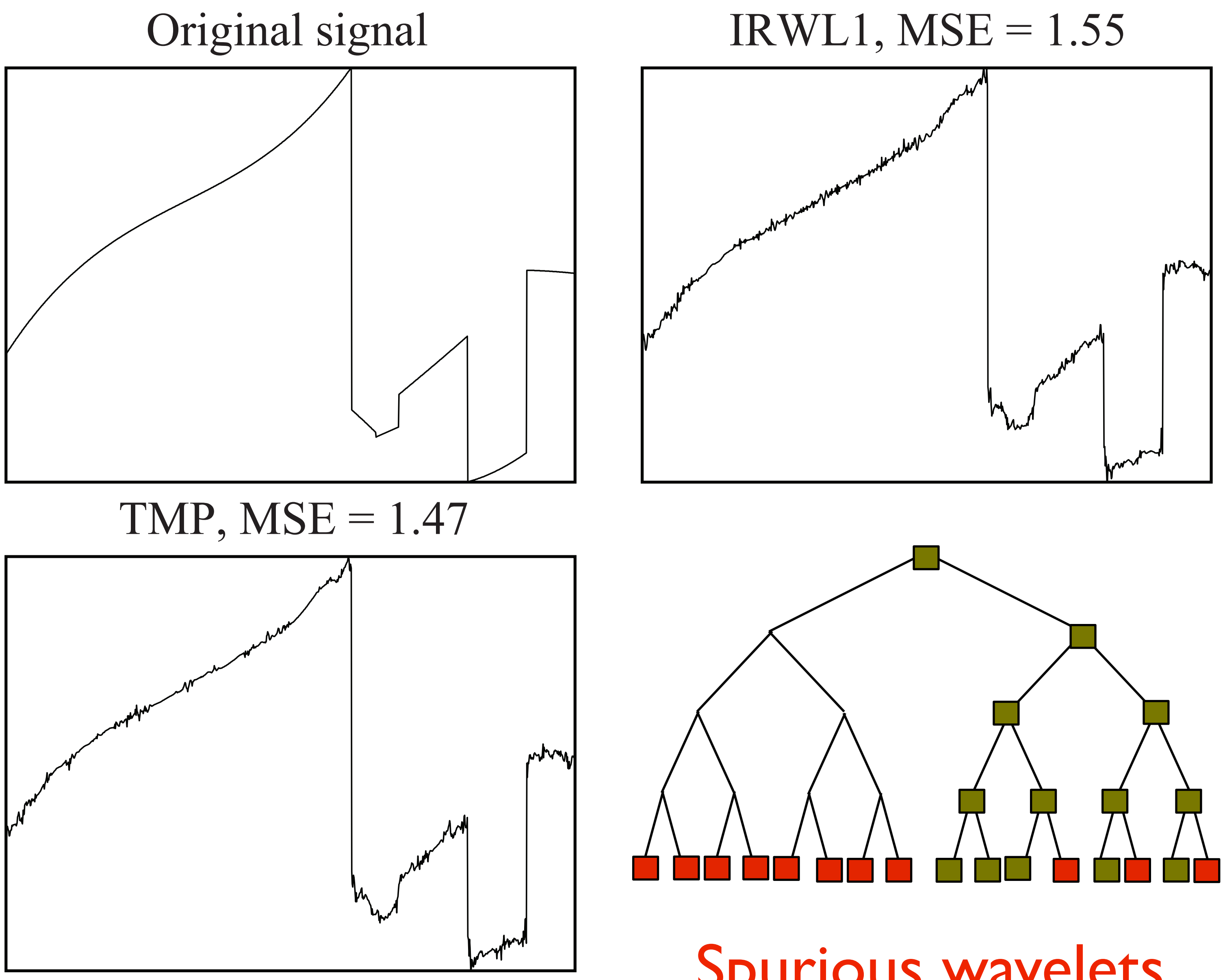
CS Reconstruction of Wavelet-Sparse Signals

$N = 1024$
$M = 300$
10 iterations/algorithm



CS Reconstruction of Wavelet-Sparse Signals

$N = 1024$
$M = 300$
10 iterations/algorithm

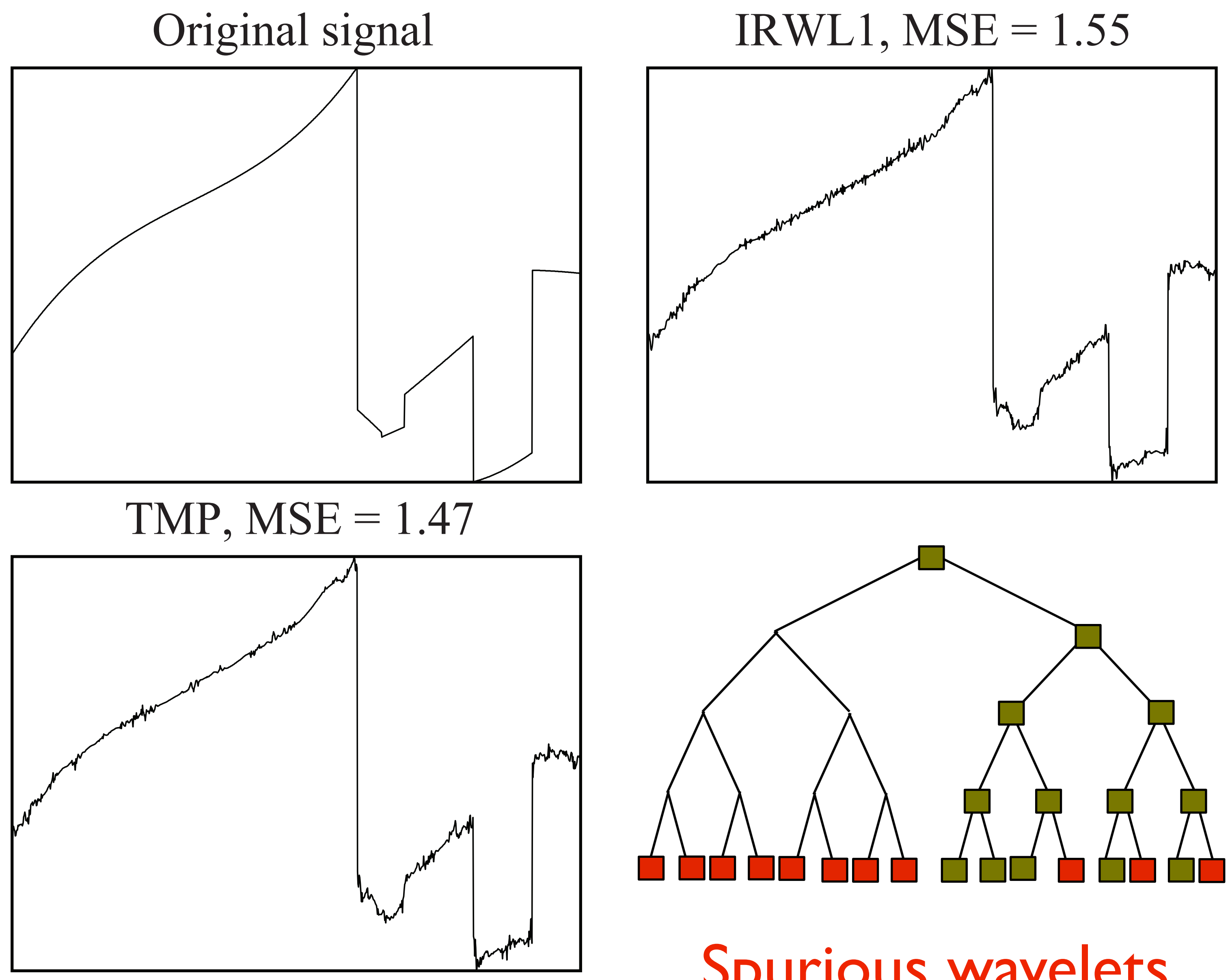


Spurious wavelets coefficients

HMT-Derived Weights for Reweighted ℓ_1 Algorithm

$N = 1024$
$M = 300$
10 iterations/algorithm

Goal for weights:
Penalize wavelet coefficients that do not follow HMT model

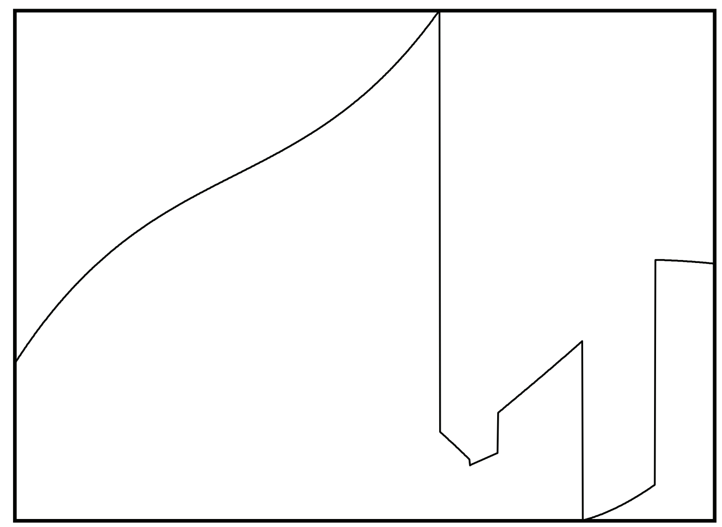


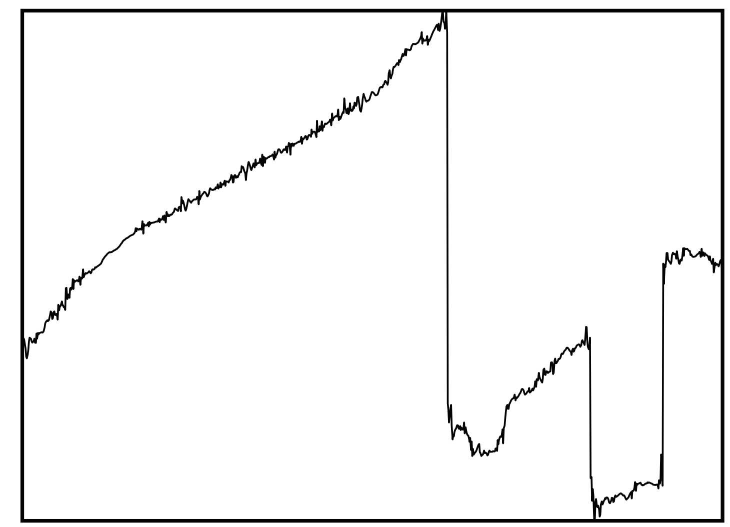
Spurious wavelets coefficients

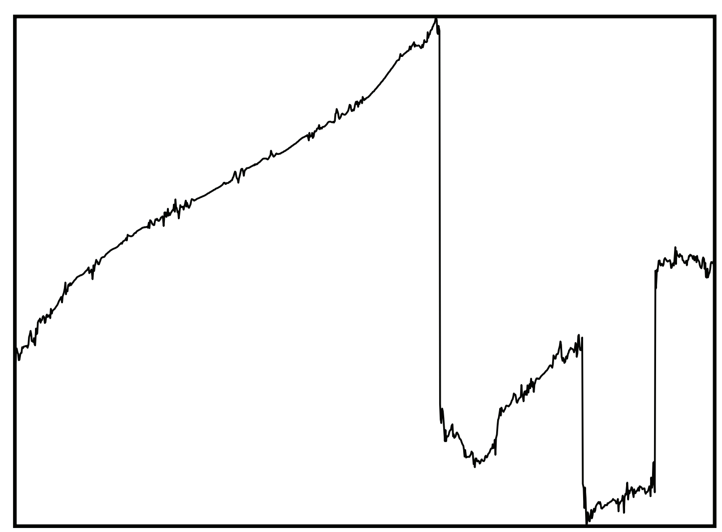
HMT-Derived Weights for Reweighted ℓ_1 Algorithm

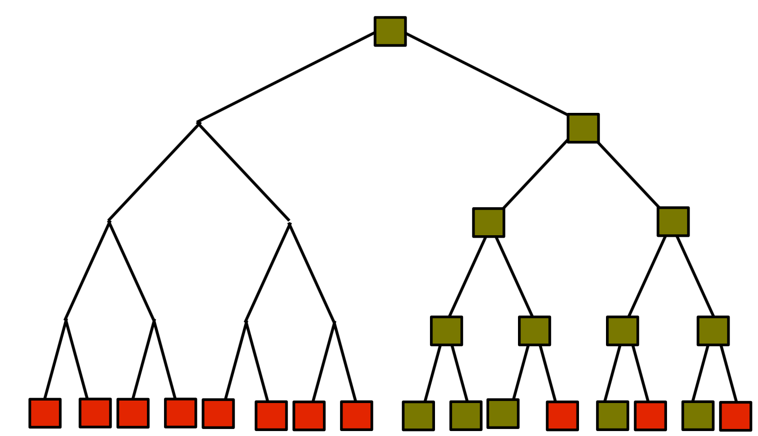
$N = 1024$
$M = 300$
10 iterations/algorithm

Goal for weights:
Penalize wavelet coefficients that do not follow HMT model







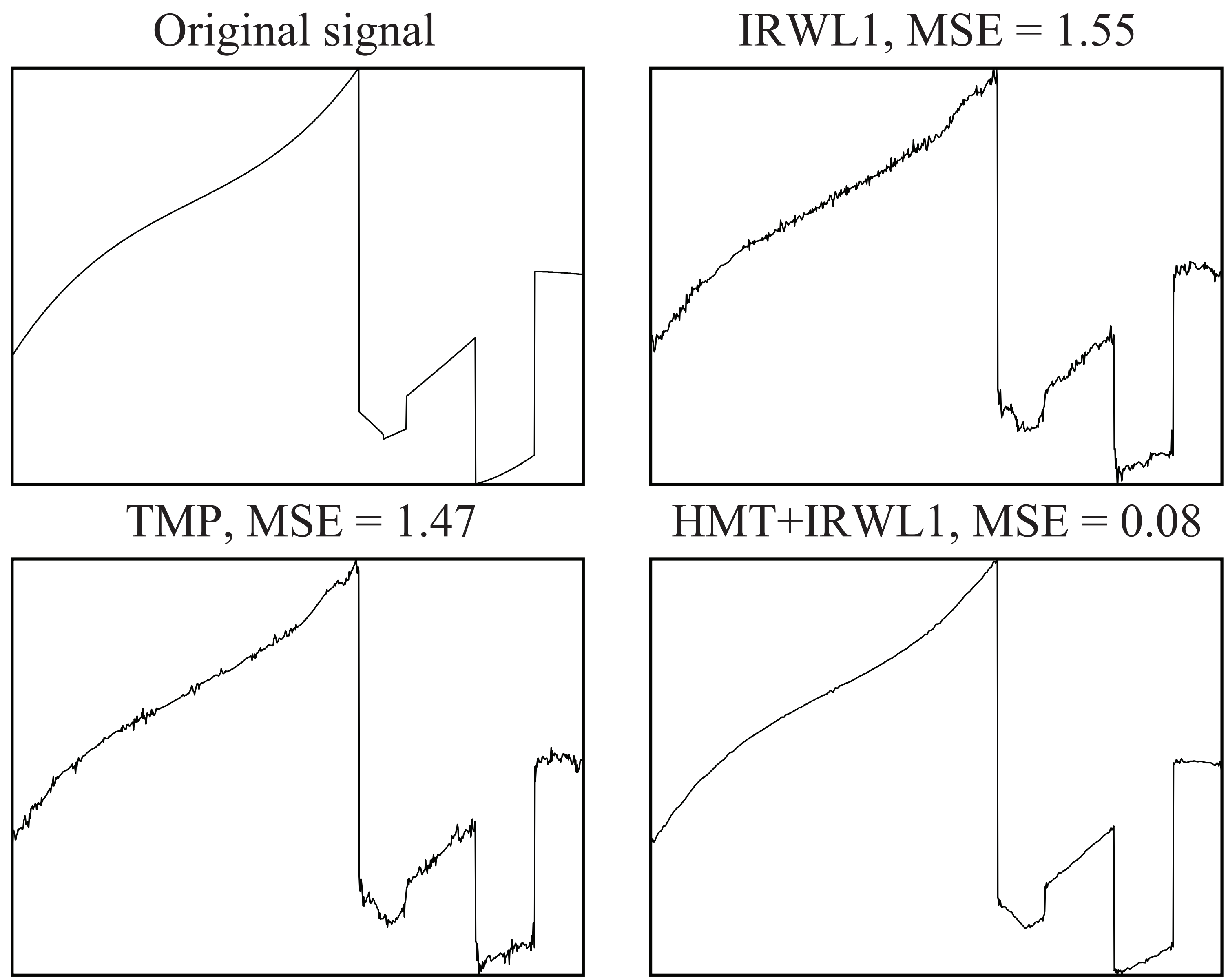


Spurious wavelets coefficients

$$w_n^{(i)} = \frac{1}{\left(p\left(S_n = \mathbf{L} | \widehat{\alpha}^{(i-1)}, \mathcal{M}\right) + \delta\right)^q} \text{ for all } n$$

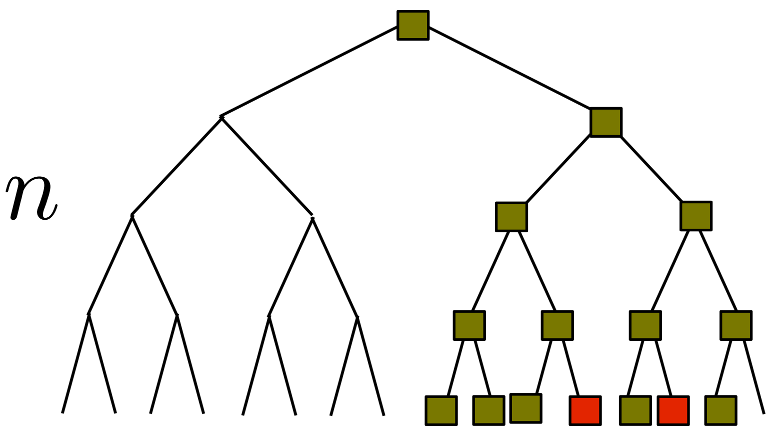
HMT-Derived Weights for Reweighted ℓ_1 Algorithm

$N = 1024$
$M = 300$
10 iterations/algorithm



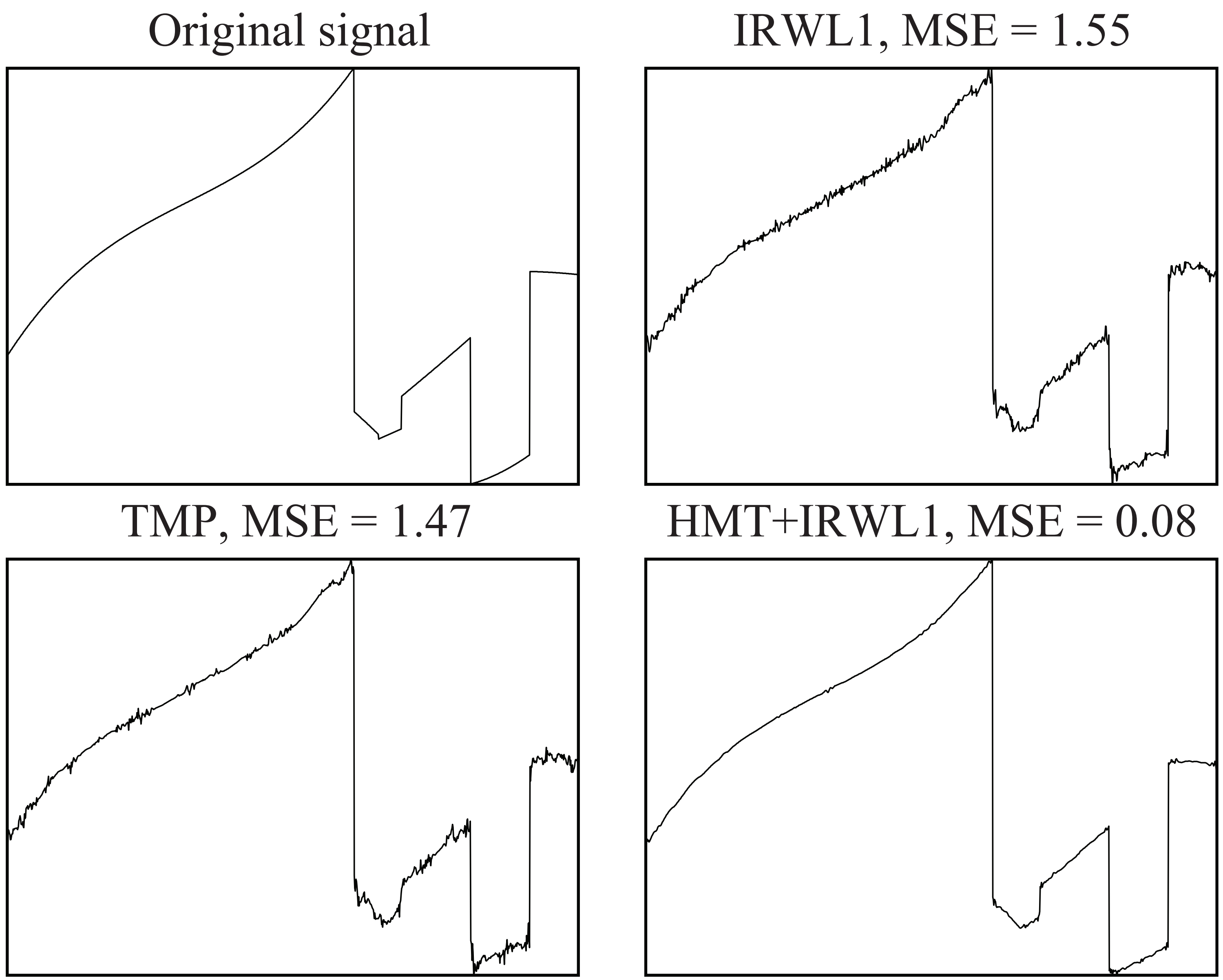
Goal for weights:
Penalize wavelet coefficients that do not follow HMT model

$$w_n^{(i)} = \frac{1}{\left(p\left(S_n = \mathbf{L} | \widehat{\alpha}^{(i-1)}, \mathcal{M}\right) + \delta\right)^q} \text{ for all } n$$



HMT-Derived Weights for Reweighted ℓ_1 Algorithm

$N = 1024$
$M = 300$
10 iterations/algorithm

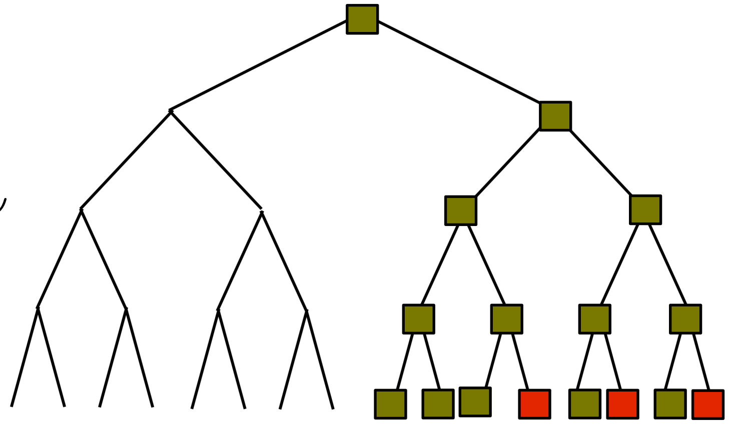


Goal for weights:
Penalize wavelet coefficients that do not follow HMT model

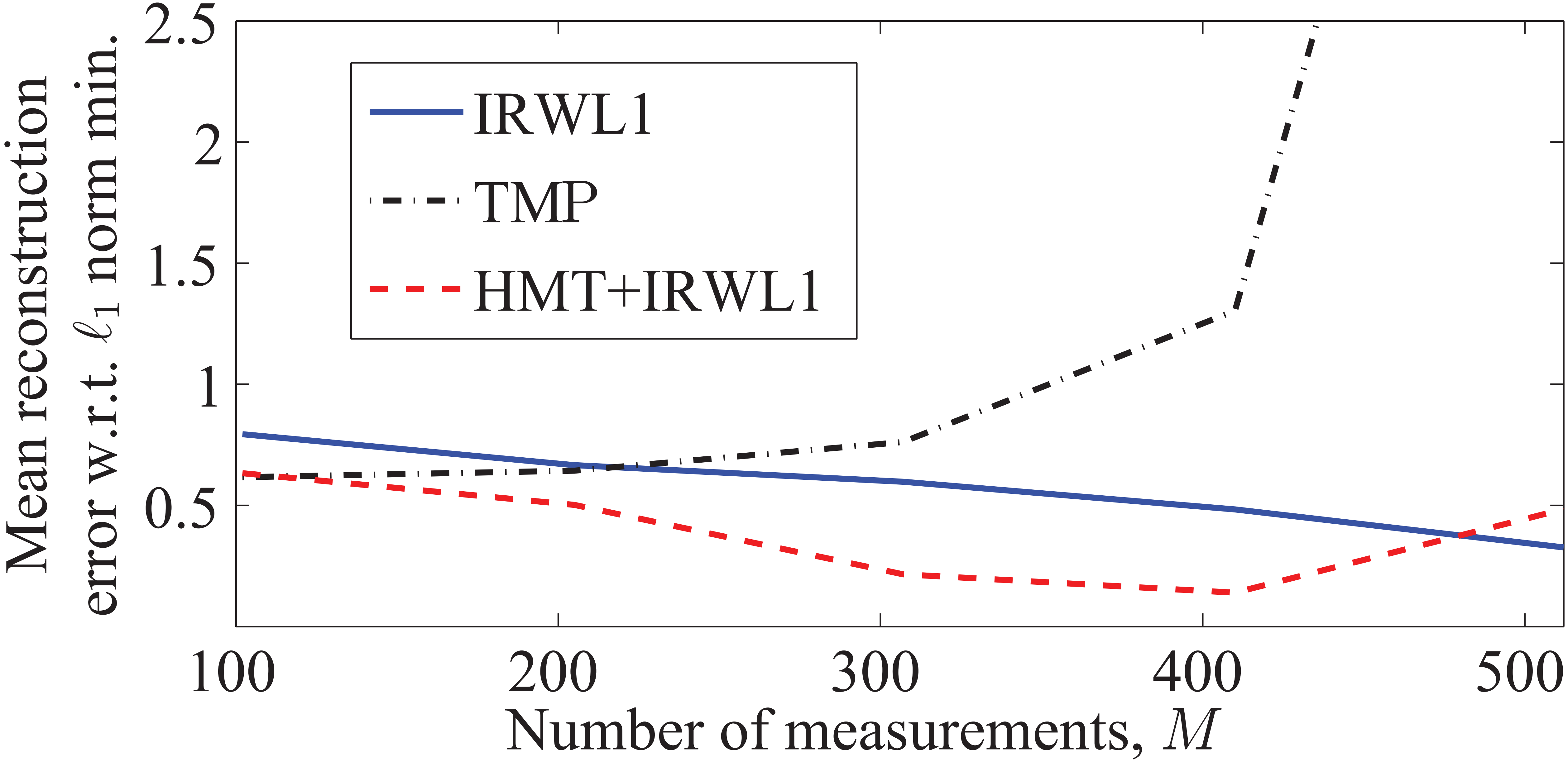
Model Training:
EM Algorithm

State Sequence Estimation:
Viterbi Algorithm

$$w_n^{(i)} = \frac{1}{\left(p\left(S_n = \mathbf{L} | \widehat{\alpha}^{(i-1)}, \mathcal{M}\right) + \delta\right)^q} \quad \text{for all } n$$



HMT-Reweighted ℓ_1 Performance



Piecewise polynomial signals
Random polynomial coefficients
5 discontinuities at random points

$N = 1024$
10 iterations for reweighting
100 repetitions for each value of M

Conclusions

- Reweighted minimization allows for signal recovery under specialized probabilistic sparse signal models
- Probability-dependent weights enforce model-fitting solutions to CS recovery
- Other sparse signal models can be used (generalized Gaussians, spatial clustering, etc.)
- Further work:
  - Analysis for reconstruction performance
  - Extensions to richer models and higher dimensions

RICE

dsp.rice.edu/cs

THE UNIVERSITY OF MICHIGAN ARTES SCIENTIA VERITAS 1817